AD-A064 810 UNIVERSITY OF SOUTHERN CALIFORNIA LOS ANGELES DEPT O--ETC F/G 12/1
MINIMUM-VARIANCE STATE ESTIMATION FOR UNIFORM CAUSAL FUNCTIONAL--ETC(U)
1978 H D WASHBURN, J M MENDEL AFOSR-75-2797
UNCLASSIFIED AFOSR-TR-79-0048 NL

1 OF 1
ADA
064810

END
DATE
FILMED
4-79
DDC

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 79-0046	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MINIMUM-VARIANCE STATE ESTIMATION FOR UNIFORM CAUSAL FUNCTIONAL EQUATIONS		5. TYPE OF REPORT & PERIOD COVERED Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) H.D. Washburn ~~and J.~~ Jerry M. Mendel		8. CONTRACT OR GRANT NUMBER(s) AFOSR-75-2797, NSF-ENG74-02297
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Southern California, Department of Electrical Engineering, Los Angeles, CA 90007		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A1
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332		12. REPORT DATE 1978
		13. NUMBER OF PAGES 10
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 11p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

DDC RECEIVED FEB 22 1979 B

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Recently, a new class of time-domain state space models has been developed ~~(Ref. 11 and 12)~~ to describe layered media systems. When layers are uniform, the resulting state equations are referred to as uniform causal functional equations. In this paper we develop the minimum-variance state estimator for such equations. We are led to a natural form of parallel data processing for practical implemenation of our estimator.

361 560

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

per Doc AD-A064812

AFOSR-TR- 79-0046

1

Presented at 17th IEEE Conference on Decision and Control San Diego, California, January 1979. Copyright by IEEE

MINIMUM-VARIANCE STATE ESTIMATION FOR UNIFORM CAUSAL FUNCTIONAL EQUATIONS

H. D. Washburn
Hughes Aircraft Company
Culver City, California 90230

J. M. Mendel
Department of Electrical Engineering
University of Southern California
Los Angeles, California 90007

Abstract

Recently, a new class of time-domain state space models has been developed (Ref. 11 and 12) to describe layered media systems. When layers are uniform, the resulting state equations are referred to as uniform causal functional equations. In this paper we develop the minimum-variance state estimator for such equations. We are led to a natural form of parallel data processing for practical implementation of our estimator.

1. Introduction

During the past few years there has been great interest in modeling layered media systems (Refs. 1-12, for example) in such diverse areas as reflection seismology, transmission lines, speech processing, optical thin coatings and EM problems. Traditionally, the models have been transfer function models; however, recently Nahi and Mendel (Ref. 11) and Mendel, et al. (Ref. 12) have shown that such systems can be modeled in the time-domain by what appears to be a new class of state equations, causal functional equations. These equations arise in a natural way when modeling lossless layered media which are described by the wave equation and boundary conditions, through the use of ray theory.

Causal functional equations are, in general, linear continuous-time equations with multiple time delays (due to layers of different travel times). They do not contain integrals or derivatives; hence, they are not integral or differential equations; nor are they finite-difference equations. As is the case with delay-time systems, they require initial value information over initial intervals of time. Because of their pure delay-nature, their impulse response is comprised of an infinite sequence of non-uniformly spaced impulse functions [see Mendel, et al. (Ref. 12) for specific details into the nature of causal functional equation].

In this paper we direct our attention at the special, but very important and useful, case of uniform causal functional equations, in which all time delays are equal. This occurs when all layers have equal travel times (layers of unequal travel time can be built up from layers of equal travel time by use of interfaces with zero reflection coefficients). We shall develop a minimum-variance state estimator for uniform causal functional equations.

Our uniform causal functional state space model, denoted $\mathcal{S}$ in the sequel, is described by the following state and observation equations:

$$\mathcal{S}\begin{cases}\underline{x}(t+\tau) = A\,\underline{x}(t) + B\,\underline{m}(t) + \underline{w}(t) & (1)\\[2ex] \underline{y}(t) = H\,\underline{x}(t) + \underline{n}(t) & (2)\end{cases}$$

In $\mathcal{S}$, $\underline{x}\varepsilon R^n$, $\underline{m}\varepsilon R^r$ is a known input, $\underline{w}\varepsilon R^n$ is a random disturbance, $\underline{y}\varepsilon R^s$, $\underline{n}\varepsilon R^s$ is random measurement noise, $A\varepsilon R^{n\times n}$, $B\varepsilon R^{n\times r}$, and $H\varepsilon R^{s\times n}$. For $\mathcal{S}$, the following initial interval information is assumed known:

$$\underline{x}(\sigma) \text{ where } \sigma \,\varepsilon\, \mathcal{J} \quad \text{and} \quad \mathcal{J} = [0,\tau) \qquad (3)$$

In Eqs. (1) and (3), τ is the uniform time delay. An example which illustrates the genesis of Eqs. (1) and (2) for seismic waves in layered media is given in Reference 11.

Some discussion is in order regarding noise processes $\underline{w}(t)$ and $\underline{n}(t)$. For differential systems we usually assume that $\underline{w}(t)$ and $\underline{n}(t)$ are white noise processes, in which case $E\{\underline{w}(t)\ \underline{w}'(\xi)\} = Q\ \delta(t-\xi)$ and $E\{\underline{n}(t)\ \underline{n}'(\xi)\} = R\ \delta(t-\xi)$. For such systems, the correlation value must contribute finite values in infinitesimal intervals. In order to have nonzero measure, the correlation value must be infinite over the infinitesimal interval. This problem does not arise for causal functional equations. In fact, we would not want $\underline{w}(t)$, for example, to be a white noise process; for, then $\underline{x}(t)$ would also be white noise, which does not make sense.

We shall show, in Theorem 1 below, that causal functional equations have solutions that are quite similar to the solutions of finite-difference equations. Our second-order noise statistics should therefore be closer to those of a discrete-time system. [Recall that if $\underline{w}(k)$ and $\underline{n}(k)$ are discrete-time white noise sequences, then $E\{\underline{w}(i)\ \underline{w}'(j)\} = Q\delta_{ij}$ and $E\{\underline{n}(i)\ \underline{n}'(j)\} = R\delta_{ij}$ where δ_{ij} is the Kronecker delta function, which is equal to unity for $i=j$ and is equal to zero for $i\neq j$.]

In our work we assume that $\underline{w}(t)$ and $\underline{n}(t)$ are mutually uncorrelated Ornstein-Uhlenbeck processes (Ref. 16 and 17). According to Parzen (Ref. 18), a stochastic process $\{\chi(t), t\geq 0\}$ is said to be an Ornstein-Uhlenbeck process with parameters $\alpha>0$ and $\beta>0$ if it is a gaussian process satisfying $E\{\chi(t)\} = 0$ and $E\{\chi(t)\ \chi(\xi)\} = \alpha\ e^{-\beta|\xi-t|}$. Feller (Ref. 19) states that the Ornstein-Uhlenbeck process is obtained by subjecting the particles of a Brownian motion to an elastic force. Parzen states further that the Ornstein-Uhlenbeck process is a model for Brownian motion which is somewhat more realistic than a Wiener process.

We can achieve the effect of a Kronecker delta in the second-order statistics of the Ornstein-Uhlenbeck process by letting $\beta\to\infty$. In this case, $E\{\chi(t)\ \chi(\xi)\} = \alpha$ if $t = \xi$ and $E\{\chi(t)\ \chi(\xi)\} = 0$ if $t \neq \xi$. More succinctly, for an Ornstein-Uhlenbeck process $\{\chi(t), t\geq 0\}$ for which $\beta\to\infty$, $E\{\chi(t)\ \chi(\xi)\} = \alpha\ \delta_{t\xi}$. We shall assume that $\underline{w}(t)$ and $\underline{n}(t)$ are vector Ornstein-Uhlenbeck processes, each of whose components has the preceding property; i.e., we assume that $\underline{w}(t)$ and $\underline{n}(t)$ are gaussian processes, for which

$$E\{\underline{w}(t)\} = \underline{0} \text{ and } E\{\underline{n}(t)\} = \underline{0} \qquad \forall\ t\epsilon\mathcal{R} \triangleq [0,\infty) \tag{4}$$

$$E\{\underline{w}(t)\ \underline{w}'(\xi)\} = Q\delta_{t\xi} \qquad \forall\ t,\ \xi\epsilon\mathcal{R} \tag{5}$$

$$E\{\underline{n}(t)\ n'(\xi)\} = R\delta_{t\xi} \qquad \forall\ t,\ \xi\epsilon\mathcal{R} \tag{6}$$

and

$$E\{\underline{w}(t)\ \underline{n}'(\xi)\} = 0 \qquad \forall\ t,\ \xi\epsilon\mathcal{R} \tag{7}$$

To complete the description of $\mathcal{S}$ we must specify the statistics of our initial interval information in (3). We shall assume that $\underline{x}(\sigma)$ $(\sigma\epsilon\mathcal{J})$ is also an Ornstein-Uhlenbeck process for which $\beta\to\infty$, and that $\underline{x}(\sigma)$ is uncorrelated with $\underline{w}(t)$ and $\underline{n}(t)$; i.e., we assume that $\underline{x}(\sigma)$ $(\sigma\epsilon\mathcal{J})$ is gaussian, and

$$E\{\underline{x}(\sigma)\} = \underline{0} \qquad \forall\ \sigma\epsilon\mathcal{J} \tag{8}$$

$$E\{\underline{x}(t)\ \underline{x}'(\xi)\} = \Lambda\delta_{t\xi} \qquad \forall\ t,\ \xi\epsilon\mathcal{J} \tag{9}$$

and

$$E\{\underline{x}(t)\ \underline{w}'(\xi)\} = 0 \text{ and } E\{\underline{x}(t)\ \underline{n}'(\xi)\} = 0 \qquad \forall\ t\epsilon\mathcal{J} \text{ and } \xi\epsilon\mathcal{R} \tag{10}$$

In many layered-media systems (e.g., a layered earth system) $\underline{x}(\sigma) = \underline{0}$, $\forall\ \sigma\epsilon\mathcal{J}$. For those systems $\Lambda = 0$ and the two conditions in Eq. (10) are satisfied because $\underline{x}(\sigma) = \underline{0}$, $\forall\ \sigma\epsilon\mathcal{J}$. For the sake of generality, we present our results below for arbitrary $\underline{x}(\sigma)$, $\sigma\epsilon\mathcal{J}$. For the sake of simplicity, all our results are given for time-invariant and stationary systems. It is straightforward to state them for time-varying and nonstationary systems [i.e., for the case when $A = A(t)$, $B = B(t)$, $H = H(t)$, $Q = Q(t)$, $R = R(t)$, and $\Lambda = \Lambda(t)$].

As we have already mentioned, system $\mathcal{S}$ is a continuous-time system that is closely related to a discrete-time formulation. If, for example, we set $t = k\tau$ in $\mathcal{S}$, where $k = 0, 1, 2, \ldots$, and we assume that $\underline{m}(t)$ and $\underline{w}(t)$ are only available at sample points $t = 0, \tau, 2\tau, \ldots$, then $\mathcal{S}$ reduces to an equivalent discrete-time system. Usually, however, τ is much larger than real data sampling rates (in geophysical applications, data is commonly sampled at 1 or 2 msec rates); hence, it would be necessary to insert many small-layers, whose interfaces have zero reflection coefficients, to convert $\mathcal{S}$ into a practical discrete-time system. That system would be of very large dimension, and it is doubtful that a Kalman filter for this high-order system would be practical. For example, if $\tau = 20$ msec and datais sampled every 1 msec, each layer (which is described by 2 states, an upgoing and a downgoing state) would be replaced by 20 layers that would be described by $2 \times 20 = 40$ states. A 100 layer system would be described by 4,000 states. By our approach, that same 100 layer system would be described by 200 states.

Section 2 presents some important preliminary results pertaining to the solution of Eq. (1) and the statistics of $\underline{x}(t)$ and $\underline{y}(t)$. A minimum-variance state estimator is derived in Section 3 first for the case when $B = 0$ and then for the more general case when $B \neq 0$. Discussions on an implementation for our state estimator are given in Section 4.

2. Preliminary Results

In this section we present results pertaining to the solution of Eq. (1) and the statistics of $\underline{x}(t)$ and $\underline{y}(t)$. Throughout this section and the rest of the paper we use the fact that $t\epsilon\mathcal{R}$ can be uniquely characterized by the mapping

$$t = t' + M\tau \quad \text{where } t'\epsilon\mathcal{J} \text{ and } M \text{ is an integer} \tag{11}$$

We depict this mapping in Figure 1.

<u>Theorem 1</u>. The solution to the uniform causal functional equation*

$$\underline{x}(t+\tau) = A\ \underline{x}(t) + \underline{w}(t)\ ;\quad x(\sigma)\quad \sigma\epsilon\mathcal{J} \tag{12}$$

is

$$\underline{x}[t' + (k+1)\tau] = A^{k+1}\ \underline{x}(t') + \sum_{i=0}^{k} A^{k-i}\ \underline{w}(t'+i\tau) \tag{13}$$

where $t'\epsilon\mathcal{J}$, $k = 0, 1, 2, \ldots$, and $t = t' + (k+1)\tau$.

Observe that (13) explicitly shows how the state at any time $t = t' + (k+1)\tau$ depends on an initial condition $\underline{x}(t')$ and the input $\underline{w}$. It is of interest to note that $\underline{x}(t)$ depends only on a <u>single element</u> of the initial values $\underline{x}(\sigma)$ $(\sigma\epsilon\mathcal{J})$

*Our preliminary results are given in this section for the $B = 0$ case. Their extensions to the $B \neq 0$ case are not needed for our Section 3 results.

namely $\underline{x}(t')$, and a finite number of point values of $\underline{w}$. This shows that the solution to the uniform causal functional state equation, although continuous-time in nature, derives its values in a discrete-time fashion for a given fixed value of $t' \in \mathcal{T}$. Of course, there are an uncountable number of points in $\mathcal{T}$; hence, we can imagine $\underline{x}(t)$ as being generated by an uncountable number of discrete-time systems.

Proof of Theorem 1: We use an iterative argument to justify the solution of Eq. (12), given in Eq. (13). First pick a $t' \in \mathcal{T}$. From Eq. (12), we see that

$$\underline{x}(t' + \tau) = A\,\underline{x}(t') + \underline{w}(t') \tag{14}$$

Again, from Eq. (12), we see that

$$\underline{x}(t' + 2\tau) = A\,\underline{x}(t' + \tau) + \underline{w}(t' + \tau) \tag{15}$$

Substitute Eq. (14) into Eq. (15) to show that

$$\underline{x}(t' + 2\tau) = A^2\underline{x}(t') + Aw(t') + \underline{w}(t' + \tau) \tag{16}$$

Iterating Eq. (12) in this same manner k+1 times, we obtain the solution form in Eq. (13).

Corollary 1. Both $\underline{x}(t)$ and $\underline{y}(t)$, $t \in \mathcal{R}$, are zero mean gaussian process.

Proof: By assumption, $\underline{x}(\sigma)$ ($\sigma \in \mathcal{T}$) and $\underline{w}(t)$ ($t \in \mathcal{R}$) are gaussian processes. Equation (13) demonstrates that $\underline{x}(t)$ ($t \in \mathcal{R}$) can be decomposed into a linear combination of these processes; hence, since linear combinations of gaussian processes are also gaussian (Ref. 13), $\underline{x}(t)$ is gaussian $\forall\, t \in \mathcal{R}$. Since $\underline{x}(\sigma)$ ($\sigma \in \mathcal{T}$) and $\underline{w}(t)$ ($t \in \mathcal{R}$) are by assumption zero mean, it follows, from Eq. (13) that $\underline{x}(t)$ ($t \in \mathcal{R}$) is also zero mean.

Because the proof of the results for $\underline{y}(t)$ is so similar to that just given for $\underline{x}(t)$, we leave it to the reader.

Next, we present some important cross-covariance results.

Theorem 2. Let t_1 and t_2 be points in . Then

$$E\{\underline{x}(t_1)\,\underline{x}'(t_2)\} = 0 \tag{17}$$

if

$$|t_1 - t_2| \neq k\tau \qquad k = 0,1,2,\ldots \tag{18}$$

This theorem states that $\underline{x}(t)$ is uncorrelated at noninteger multiples of the uniform delay time τ.

Proof: Time points t_1 and t_2 can be expressed by means of Eq. (11) as

$$t_1 = t' + M\tau \qquad t' \in \mathcal{T} \text{ and } M \text{ an integer} \tag{19a}$$

$$t_2 = t'' + J\tau \qquad t'' \in \mathcal{T} \text{ and } J \text{ an integer} \tag{19b}$$

From Eqs. (19a), (19b) and (13), we find that

$$\begin{aligned} E\{\underline{x}(t_1)\underline{x}'(t_2)\} &= E\{\underline{x}(t'+M\tau)\underline{x}'(t''+J\tau)\} \\ &= E\{[A^M\underline{x}(t') + \sum_{i=0}^{M-1} A^{M-1-i}\underline{w}(t'+i\tau)] \\ &\quad [A^J\underline{x}(t'') + \sum_{j=0}^{J-1} A^{J-1-j}\,\underline{w}(t''+j\tau)]'\} \end{aligned} \tag{20}$$

If $t' \neq t''$, then the right-hand side of Eq. (20) evaluates to zero, because of Eqs. (5), (9), and (10). We must now ascertain when $t' \neq t''$.

We shall show that

$$t' \neq t'' \quad \text{iff} \quad |t_1 - t_2| = k\tau \qquad k = 0,1,2,\ldots \tag{21}$$

Our proofs of both the necessity and sufficiency of Eq. (21) are by the method of contradiction. To begin, we assume the truth of $t' \neq t''$ and assume also that $t_1 - t_2 = P\tau$, where $P = \ldots,-2,-1,0,1,2,\ldots$. From Eqs. (19a) and (19b), this means that

$$t' - t'' + (M - J)\tau = P\tau \tag{22}$$

or that

$$t' - t'' = (P - M + J)\tau \tag{23}$$

Now $t', t'' \in \mathcal{T} = [0,\tau)$; hence,

$$-\tau < t' - t'' < \tau \tag{24}$$

Consequently, since P, M, and J are all integers, the only way Eq. (23) can be true is if $P-M+J = 0$, *in which case* $t' = t''$; *but, this contradicts the* assumption that $t' \neq t''$. We conclude, therefore, that if $t' \neq t''$, then $t_1 - t_2 \neq P\tau$, $P = \ldots,-2,-1,0,1,2,\ldots$, which can also be stated as $|t_1-t_2| \neq k\tau$, $k = 0,1,2,\ldots$.

Next, we assume the truth of $|t_1-t_2| \neq k\tau$, $k = 0,1,2,\ldots$, and assume also that $t' = t''$. From Eqs. (19a) and (19b), this means that

$$t_1 - t_2 = (M - J)\tau \quad ; \tag{25}$$

but, this contradicts our assumption that $|t_1-t_2| \neq k\tau$. We conclude, therefore, that if $|t_1-t_2| \neq k\tau$, then $t' \neq t''$. This completes the proof of Thoerem 2.

Corollary 2. Let t_1 and t_2 be points in $\mathcal{R}$. Then

$$E\{\underline{y}(t_1)\underline{y}'(t_2)\} = 0 \tag{26}$$

and

$$E\{\underline{y}(t_2)\underline{x}'(t_1)\} = 0 \ , \tag{27}$$

if

$$|t_1 - t_2| \neq k\tau \qquad k = 0,1,2,\ldots \tag{28}$$

Proof: From Eq. (2), we see that

$$E\{\underline{y}(t_1)\underline{y}'(t_2)\} = HE\{\underline{x}(t_1)\underline{x}'(t_2)\}H' + E\{\underline{n}(t_1)\underline{n}'(t_2)\} + HE\{\underline{x}(t_1)\underline{n}'(t_2)\} + E\{\underline{n}(t_1)\underline{x}'(t_2)\}H' \quad (29)$$

The first term in this expression is zero via Theorem 2. Observe that Eq. (28) implies that $t_1 \neq t_2$; hence, the last three terms in Eq. (29) are zero via Eqs. (16), (13) and (10). The proof of Eq. (27) follows in a similar manner.

In Section 3, where we derive the minimum-variance estimator for $\underline{x}(t)$, we assume that measurements, $\underline{y}(t)$, are available from time zero up to and including time t. We now introduce three data sets. Let Y(t) be the set of measurement data which are available up to and including time t; i.e.,

$$Y(t) = \{y(\lambda): \; 0 \leq \lambda \leq t, \; t \epsilon \mathcal{R}\} \quad (30)$$

Let $Y_M(t')$ be a finite set of measurements associated with point $t' \epsilon \mathcal{J}$ and integral multiples of τ from t'; i.e.,

$$Y_M(t') = \{\underline{y}(t'+j\tau), \; j = 0,1,2,\ldots,M\} \quad (31)$$

For j = 0, $Y_M(t')$ only contains the single measurement $\underline{y}(t')$. For j = M, $Y_M(t')$ contains M+1 measurements made at the points labeled t'. $t'+\tau$, $\ldots, t'+M\tau$ in Figure 1. Finally, let $Y_C(t')$ be the set difference between Y(t) and $Y_M(t')$; i.e.,

$$Y_C(t') = Y(t) - Y_M(t') \quad (32)$$

Observe from Eqs. (30), (31), and (32) that

$$Y(t) = Y_M(t') \cup Y_C(t') \quad (33)$$

and

$$Y_C(t') \cap Y_M(t') = \phi \quad (34)$$

where ϕ is the null set.

Theorem 3. Let t be a point in $\mathcal{R}$ such that the representation in Eq. (11) is valid. Then $Y_M(t')$ is statistically independent of $Y_C(t')$ and $\underline{x}(t)$ is statistically independent of $Y_C(t')$.

Proof: We represent t as in Eq. (11) and choose a typical element, $\underline{y}(t_2)$, from data set $Y_C(t')$; i.e.,

$$\underline{y}(t_2) \epsilon Y_C(t') \quad (35)$$

By construction of $Y_C(t')$ we know that

$$t_2 = t'' + j\tau \qquad j = 0,1,\ldots,M \quad (36)$$

and $t'' \epsilon \mathcal{J}$. We observe that $t'' \neq t'$; for, if $t' = t''$ then $t_2 = t' + j\tau$ $(j = 0,1,\ldots,M)$, and these time points are associated with $Y_M(t')$. From our preceding discussions, we know that $Y_M(t')$ and $Y_C(t')$ share no common time points; hence, $t'' \neq t'$.

Since $t' \neq t''$, we conclude, from Eq. (21), that

$$|t - t_2| \neq k\tau \qquad k = 0,1,\ldots \quad (37)$$

which means that time points in $Y_C(t')$ satisfy Eq. (28); hence, from Eq. (27), we see that

$$E\{\underline{y}(t_2)\underline{x}'(t)\} = 0 \quad \text{for} \; \forall \; \underline{y}(t_2) \epsilon Y_C(t') \quad (38)$$

Since $\underline{x}(t)$ and $\underline{y}(t)$ are gaussian (Corollary 1), the fact in Eq. (38), that $\underline{y}(t_2)$ and $\underline{x}(t)$ are uncorrelated for $\forall \; \underline{y}(t_2) \epsilon Y_C(t')$ also means that $\underline{y}(t_2)$ and $\underline{x}(t)$ are statistically independent (Ref. 13) for $\forall \; \underline{y}(t_2) \epsilon Y_C(t')$.

Next, pick a typical element, $\underline{y}(t_1)$, from $Y_M(t')$; i.e.,

$$\underline{y}(t_1) \epsilon Y_M(t') \quad (39)$$

By construction of $Y_M(t')$, we know that

$$t_1 = t' + j\tau \qquad j = 0,1,2,\ldots,M \quad (40)$$

Since $t' \neq t''$, we know that t_1 and t_2 satisfy Eq. (28); hence, it follows from Eq. (26) that

$$E\{\underline{y}(t_1)\underline{y}'(t_2)\} = 0 \quad \forall \; \underline{y}(t_2) \epsilon Y_C(t') \text{ and } \underline{y}(t_1) \epsilon Y_M(t') \quad (41)$$

From Corollary 1, this once again means that $\underline{y}(t_1)$ and $\underline{y}(t_2)$ are statistically independent for $\forall \; \underline{y}(t_2 \epsilon Y_C(t')$ and $\underline{y}(t_1) \epsilon Y_M(t')$; or, that $Y_C(t')$ and $Y_M(t')$ are statistically independent. This completes the proof of Theorem 3.

3. Minimum-Variance State Estimator

To begin, we present the minimum-variance state estimator for system $\mathcal{S}$ with B=0. The case of a known forcing function in the state equation is treated below in Theorem 5.

Theorem 4. For system $\mathcal{S}$, with B=0, the minimum-variance estimator of $\underline{x}(t)$, where $t = t' + M\tau$, $t' \epsilon \mathcal{J}$ and M an integer, is

$$\hat{\underline{x}}(t) = E\{\underline{x}(t)|Y_M(t')\} \quad (42)$$

This theorem states the interesting result that the optimal estimator of $\underline{x}(t)$ need not be conditioned on the entire data set, Y(t). The optimal estimator of $\underline{x}(t)$ need only be conditioned on data set $Y_M(t')$, a data set which contains only a finite number of points.

Proof: From estimation theory (Refs. 14 and 15, for example), it is well known that the minimum variance estimator of x(t) is

$$\hat{\underline{x}}(t) = E\{\underline{x}(t)|Y(t)\} \quad (43)$$

which can also be written, using Eq. (33), as

$$\hat{\underline{x}}(t) = E\{\underline{x}(t)|Y_M(t'), \; Y_C(t')\} \quad (44)$$

Next, we use the fact (proven in Meditch (Ref. 14), pg. 101) that if $\underline{a}$, $\underline{b}$, and $\underline{c}$ are gaussian random vectors and $\underline{b}$ and $\underline{c}$ are statistically independent, then

$$E\{\underline{a}|\underline{b},\underline{c}\} = E\{\underline{a}|\underline{b}\} + E\{\underline{a}|\underline{c}\} - E\{\underline{a}\} \tag{45}$$

Since $\underline{x}(t)$, $Y_M(t')$ and $Y_C(t')$ are gaussian and $Y_M(t')$ and $Y_C(t')$ are statistically independent (Theorem 3), we can expand the right-hand side of Eq. (44) by means of Eq. (45), to obtain

$$\hat{\underline{x}}(t) = E\{\underline{x}(t)|Y_M(t')\} + E\{\underline{x}(t)|Y_C(t')\} - E\{\underline{x}(t)\} \tag{46}$$

Additionally, since $\underline{x}(t)$ and $Y_C(t')$ are statistically independent (Theorem 3), $E\{\underline{x}(t)|Y_C(t')\} = E\{\underline{x}(t)\}$; thus, Eq. (46) reduces to Eq. (42).

Corollary 3. For any fixed $t' \in \mathcal{J}$, $\hat{\underline{x}}(t) = \hat{\underline{x}}(t'+M\tau)$ can be computed by the discrete-time Kalman filtering equations with t' considered as the initial starting time. Hence, for $t' \in \mathcal{J}$, compute $\hat{\underline{x}}(t'+M\tau)$ from the following equations:

$$\hat{\underline{x}}(t'+M\tau) = A\hat{\underline{x}}[t' + (M-1)\tau] + K(t'+M\tau)\{\underline{y}(t'+M\tau) - HA\hat{\underline{x}}[t' + (M-1)\tau]\} \tag{47}$$

$$P[t'+M\tau|t' + (M-1)\tau] = AP[t' + (M-1)\tau|t' + (M-1)\tau]A' + Q \tag{48*}$$

$$K(t'+M\tau) = P[t'+M\tau|t'+(M-1)\tau] \cdot H'(HP[t'+M\tau|t'+(M-1)\tau]H' + R)^{-1} \tag{49*}$$

and

$$P(t'+M\tau|t'+M\tau) = [I-K(t'+M\tau)H] \cdot P[t'+M\tau|t'+(M-1)\tau] \tag{50}$$

where $M = 1,2,\ldots$, and, $\hat{\underline{x}}(t') = E\{\underline{x}(t')\}$ and $P(t'|t') = E\{\underline{x}(t')\underline{x}'(t')\} = \Lambda$.

Proof: Equation (42) states that $\hat{\underline{x}}(t)$ is obtained from a finite point measurement data set. It can be expressed by means of Eqs. (11) and (31) as

$$\hat{\underline{x}}(t) = E\{\underline{x}(t'+M\tau)|\underline{y}(t'),\underline{y}(t'+\tau),\ldots,\underline{y}(t'+M\tau)\} \tag{51}$$

The right-hand side of Eq. (51) is analogous to the optimal filtered estimate for a discrete-time system with sampling rate equal to τ sec and initial time equal to t'. Our notation in Eq. (47) for $\hat{\underline{x}}(t'+M\tau)$ is equivalent, therefore, to the notation $\hat{\underline{x}}(t'+M\tau|t'+M\tau)$. Equations (47) through (50) are the discrete-time Kalman filter equations (Ref. 14, for example) obtained by replacing discrete variable k in those equations by $t' + M\tau$.

*For nonstationary noise process, Q becomes $Q[t' + (M-1)\tau]$ and R becomes $R(t' + M\tau)$.

We observe, from Corollary 3, that for each fixed $t' \in \mathcal{J}$, the optimal estimate of the state at $t' + M\tau$ ($M = 1,2,\ldots$) is obtained by iterating the familiar Kalman filter equations on τ, using t' as the initial starting time. Since there are an uncountable number of points in $\mathcal{J}$, this results in an uncountable number of estimators which have to be implemented in order to obtain $\hat{\underline{x}}(t)$ for any $t \in \mathcal{R}$. We discuss a practical implementation of these results in Section IV.C.

Next, we shall extend our Theorem 4 and Collary 3 results to the case where there is a known forcing fucntion in the state equation, that is, to the case when $B \neq 0$ in .

Theorem 5. Let $\hat{\underline{x}}(t)$ denote the minimum-variance estimator of $\underline{x}(t)$ for $\mathcal{A}$ in Eqs. (1) and (2). Let $\hat{\underline{x}}_1(t)$ denote the minimum-variance estimator of $\underline{x}(t)$ under the assumption that B=0 in Eq. (1). Additionally, $\hat{\underline{x}}_1(t)$ is associated with the following modified measurement equation:

$$\underline{y}_1(t) = \underline{y}(t) - H\,\underline{g}(t) \tag{52}$$

where $\underline{g}(t) \in \mathcal{R}^n$ satisfies the uniform causal functional equation

$$\underline{g}(t+\tau) = A\,\underline{g}(t) + B\,\underline{m}(t) \tag{53}$$

for which

$$\underline{g}(\sigma) = 0 \qquad \forall\, \sigma \in \mathcal{J}. \tag{54}$$

Then,

$$\hat{\underline{x}}(t) = \hat{\underline{x}}_1(t) + \underline{g}(t) \quad . \tag{55}$$

The proof of this theorem is given in Appendix A. The essence of the theorem is that $\hat{\underline{x}}_1(t)$ can be computed using Theorem 4 or Corollary 3 for the modified measurement $\underline{y}_1(t)$ which can be constructed a priori from $\underline{y}(t)$ and $\underline{g}(t)$. Signal $\underline{g}(t)$ can be precomputed since $\underline{m}(t)$ is known a priori. The solution of Eqs. (53) and (54) is given by Eq. (13) where $\underline{x}(t') = \underline{0}$ and $\underline{w}$ is replaced by $B\underline{m}$. We then obtain the desired estimate of $\underline{x}(t)$, $\hat{\underline{x}}(t)$, by the simple linear transformation in Eq. (55).

Theorem 5 provides the formal results for $\hat{\underline{x}}(t)$. We do not recommend calculating $\hat{\underline{x}}(t)$ by Theorem 5; for the calculations involve solving an auxiliary causal functional equation. A more practical way for computing $\hat{\underline{x}}(t)$ is given in the following:

Corollary 4. Let $\hat{\underline{x}}(t)$ denote the minimum-variance estimator of $\underline{x}(t)$ for $\mathcal{A}$ in Eqs. (1) and (2). For any fixed $t' \in \mathcal{J}$, $\hat{\underline{x}}(t) = \hat{\underline{x}}(t'+M\tau)$ can be computed from the following:

$$\hat{\underline{x}}(t'+M\tau) = A\hat{\underline{x}}[t'+(M-1)\tau] + B\underline{m}[t'+(M-1)\tau] + K(t'+M\tau)\{\underline{y}(t'+M\tau) - HA\hat{\underline{x}}[t'+(M-1)\tau] - HB\underline{m}[t'+(M-1)\tau]\}\ , \tag{56}$$

where $P[t'+M\tau|t'+(M-1)\tau]$, $K(t'+M\tau)$, and

$P(t'+M\tau|t'+M\tau)$ are given by Eqs. (48), (49), and (50), respectively, and $M = 1,2,\ldots$. Additionally, $\hat{x}(t') = E\{\underline{x}(t')\}$ and $P(t'|t') = E\{\underline{x}(t')\underline{x}'(t')\} = \Lambda$.

<u>Proof</u>: For $t = t'+M\tau$, we can express $\hat{x}(t'+M\tau)$ by means of Eqs. (55), (53) and (47), as

$$\begin{aligned}\hat{\underline{x}}(t'+M\tau) &= \hat{\underline{x}}_1(t'+M\tau) + \underline{g}(t'+M\tau)\\ &= A\hat{\underline{x}}_1[t'+(M-1)\tau] + K(t'+M\tau)\{\underline{y}_1(t'+M\tau)\\ &\quad - HA\hat{\underline{x}}_1[t'+(M-1)\tau]\} + A\,\underline{g}[t'+(M-1)\tau]\\ &\quad + B\underline{m}[t'+(M-1)\tau] \end{aligned} \tag{57}$$

where, from Eqs. (52) and (53),

$$\begin{aligned}\underline{y}_1(t'+M\tau) &= \underline{y}(t'+M\tau) - H\,\underline{g}(t'+M\tau)\\ &= \underline{y}(t'+M\tau) - HA\underline{g}[t'+(M-1)\tau]\\ &\quad - HB\underline{m}[t'+(M-1)\tau]\end{aligned} \tag{58}$$

Substitute Eq. (58) into Eq. (57), making use of Eq. (55), to show that

$$\begin{aligned}\hat{\underline{x}}(t'+M\tau) &= A\hat{\underline{x}}[t'+(M-1)\tau] + B\underline{m}[t'+(M-1)\tau]\\ &\quad + K(t'+M\tau)\{\underline{y}(t'+M\tau) - HA\hat{\underline{x}}[t'+(M-1)\tau]\\ &\quad - HB\underline{m}[t'+(M-1)\tau]\}\end{aligned} \tag{59}$$

which is our desired result, Eq. (56). Observe that gain matrix $K(t'+M\tau)$ in Eq. (59) is the same gain matrix as in Eq. (47); hence, it is computed from Eq. (49), and $P[t'+M\tau|t'+(M-1)\tau]$ and $P(t'+M\tau|t'+M\tau)$ are computed from Eqs. (48) and (50), respectively.

We see that a known forcing function in a uniform causal functional state equation is handled in the estimator equation in exactly the same manner that such a function is handled for a finite-difference equation. It does not affect the error covariance equations.

4. Computation

A. <u>Introduction</u>

In Section 2, we demonstrated that the solution, $\underline{x}(t)$, of causal functional equation (1) can be generated by an uncountable number of discrete-time systems. In Section 3 we demonstrated that $\hat{\underline{x}}(t)$ ($t\in\mathcal{R}$) can be generated by means of an uncountable number of discrete-time Kalman filters. When we simulate our results on a digital computer, computations are made every T sec. at discrete time points. Consequently, on a digital computer, $\underline{x}(t)$ is generated by a countable number of discrete-time systems, and $\hat{\underline{x}}(t)$ is generated by a countable number of discrete-time Kalman filters. We explore the detailed ramifications of these observations below.

B. <u>Digital Computer Simulation of Uniform Causal Functional Equation</u>

We assume here that τ is an integral multiple of sample rate T; i.e., we assume that

$$\tau = LT \tag{60}$$

The approximation denotes a temporal quantization which may be needed to associate a value of L with τ. For $\tau = LT$, $\mathcal{J} \to \mathcal{J}_C$, where

$$\mathcal{J}_C = \{t': t' = \ell T,\ \ell = 0,1,\ldots,L-1\} \tag{61}$$

Set $\mathcal{J}_C$ has a countable number of elements. It does not include $t=\tau$, since set $\mathcal{J}$ does not include that point.

We direct our attention at Eq. (13) which is the solution to uniform causal functional equation (21). Observe from Eqs. (60) and (61) that

$$t' + p\tau = (\ell + Lp)T \tag{62}$$

Letting

$$\underline{x}_\ell(p) \triangleq \underline{x}[(\ell + Lp)T] \tag{63}$$

and

$$\underline{w}_\ell(p) \triangleq \underline{w}[(\ell + Lp)T] \quad , \tag{64}$$

We see that (13) can be written as

$$x_\ell(k+1) = A^{k+1}\,\underline{x}_\ell(0) + \sum_{i=0}^{k} A^{k-i}\,\underline{w}_\ell(i) \tag{65}$$

where $\ell = 0,1,\ldots,L-1$. This equation is the well-known solution to the following finite-difference equation (Ref. 20, for example):

$$\underline{x}_\ell(k+1) = A\,\underline{x}_\ell(k) + \underline{w}_\ell(k) \tag{66}$$

where $\ell = 0,1,\ldots,L-1$ and $k = 0,1,2,\ldots,k^*$. We conclude, therefore, that the sampled solution to our uniform causal functional equation can be generated as depicted in Figure 2.

To begin, we must fix our useful data length at an integer multiple of τ, say $N\tau$. In terms of our sampling rate, T, our useful data length is NLT sec.

On a discrete time-scale the Ornstein-Uhlenbeck process, with $\beta\to\infty$, reduces to a discrete-time white noise sequence $\{\underline{w}(jT),\ j = 0,1,\ldots,J^*\}$ where $E\{\underline{w}(jT)\ w'(iT)\} = Q\delta_{ij}$. In order to cover the useful data length, we choose

$$J^* = NL \tag{67}$$

we create sequences $\underline{w}_0(k), \underline{w}_1(k), \ldots, \underline{w}_{L-1}(k)$, which are needed to simulate Eq. (66) by sorting and distributing $\{\underline{w}(jT),\ j = 0,1,\ldots,NL\}$ according to the <u>distribution algorithm</u> given on Figure 2.

Index k ranges from zero to k^*. To determine k^*, we observe that the very last value of $\underline{w}[(\ell + Lk)T]$ used by the distribution algorithm (set ℓ = L-1 and k = k*) is $\underline{w}\{[(k^*+1)L-1]T\}$. Since $\mathcal{J}$ is an open interval on its right-hand end, the argument of this function must equal NL-1; hence,

$$k^* = N-1 \tag{68}$$

Our L noise sequences, $\underline{w}_0(k)$, $\underline{w}_1(k)$, ..., $\underline{w}_{L-1}(k)$, drive L systems, each given by Eq. (66), which operate in parallel. Solving our uniform causal functional equation on a digital computer has led to a parallel data processing algorithm.

As a final step we generate $\{\underline{x}(jT),\ j = 0, 1,\ldots,NL-1\}$ by multiplexing $\underline{x}_0(k)$, $\underline{x}_1(k)$, ..., $\underline{x}_{L-1}(k)$ using the multiplexor algorithm also given on Figure 2. It is straightforward to show that

$$\{\underline{x}(jT),\ j=0,1,\ldots,NL-1\} = \{\underline{x}_0(0),\underline{x}_1(0),\ldots,\underline{x}_{L-1}(0), \underline{x}_0(1),\underline{x}_1(1),\ldots,\underline{x}_{L-1}(1),\ldots,\underline{x}_0(k^*),\underline{x}_1(k^*), \ldots,\underline{x}_{L-1}(k^*)\} \tag{69}$$

Also, given k and ℓ, j = ℓ+Lk.

C. Digital Computer Simulation of Minimum-Variance Estimator

We now direct out attention at the simulation of Eqs. (47)-(50). As in Section B, we assume that time is discretized so that Eqs. (60) and (61) are true. Letting

$$\hat{\underline{x}}_\ell(M|M) \triangleq \hat{\underline{x}}[(\ell + LM)T] \tag{70}$$

$$K_\ell(M) \triangleq K[(\ell + LM)T] \tag{71}$$

$$P_\ell(M|M-1) \triangleq P[(\ell + LM)T|(\ell + LM - L)T] \tag{72}$$

$$P_\ell(M|M) \triangleq P[(\ell + LM)T|(\ell + LM)T] \tag{73}$$

and

$$y_\ell(M) \triangleq y[(\ell + LM)T] \ , \tag{74}$$

we rewrite Eqs. (47)-(50) as:

$$\hat{\underline{x}}_\ell(M|M) = A\hat{\underline{x}}_\ell(M-1|M-1) + K_\ell(M)[\underline{y}_\ell(M) - HA\hat{\underline{x}}_\ell(M-1|M-1)] \tag{75}$$

$$P_\ell(M|M-1) = AP_\ell(M-1|M-1)A' + Q \tag{76}$$

$$K_\ell(M) = P_\ell(M|M-1)H'[HP_\ell(M|M-1)H' + R]^{-1} \tag{77}$$

and

$$P_\ell(M|M) = [I - K_\ell(M)H]\ P_\ell(M|M-1) \tag{78}$$

where for each ℓ (ℓ=0,1,...,L-1), M = 1,2,..., M* = N-1, and,

$$\hat{\underline{x}}_\ell(0|0) = E\{\underline{x}_\ell(0)\} \tag{79}$$

$$P_\ell(0|0) = \Lambda \ . \tag{80}$$

We refer to Eqs. (75) - (80) as the ℓ^{th} Kalman filter, KF_ℓ.

We obtain $\hat{\underline{x}}(jT|jT)$ (j=1,2,...,J* = NL-1) by means of the parallel data processing algorithm depicted in Figure 3.

The Kalman filters do not generate estimates over the initial time interval $\mathcal{J}_C$; for over that interval $\hat{\underline{x}}_\ell(0|0)$ is given a priori. The first measurement used by KF_0 is $\underline{y}(LT)$, whereas the last measurement used by KF_{L-1} is $\underline{y}[(NL-1)T]$; hence, we must have the sequence $\{\underline{y}(jT),\ j=L,L+1,\ldots,NL-1\}$ available at the start of the simulation. Kalman filter outputs $\hat{\underline{x}}_0(M|M),\hat{\underline{x}}_1(M|M),\ldots,\hat{\underline{x}}_{L-1}(M|M)$ are multiplexed to give the desired estimates

$$\{\hat{\underline{x}}(jT|jT),j=L,L+1,\ldots,NL-1\} = \{\hat{\underline{x}}_0(1|1),\hat{\underline{x}}_1(1|1), \ldots,\hat{\underline{x}}_{L-1}(1|1),\hat{\underline{x}}_0(2|2),\hat{\underline{x}}_1(2|2),\ldots,\hat{\underline{x}}_{L-1}(2|2), \ldots,\hat{\underline{x}}_0(N-1|N-1),\hat{\underline{x}}_1(N-1|N-1), \ldots,\hat{\underline{x}}_{L-1}(N-1|N-1)\} \tag{81}$$

For a time-varying system $\mathcal{S}$, we must make the following substitutions in Eqs. (75)-(78): $A_\ell(M) \to A$, $H_\ell(M) \to H$, $Q_\ell(M-1) \to Q$, and $R_\ell(M) \to R$, where, for example,

$$A_\ell(M) \triangleq A[(\ell + LM)T] \tag{82}$$

In this case, each of the L Kalman filters has no common calculations, and the computational burden *can be quite heavy*.

If, on the other hand, system $\mathcal{S}$ is time-invariant or slowly time-varying (i.e., all matrices are piecewise constant over τ sec intervals), then Eqs. (76), (77), and (78) are not functions of ℓ (remember that ℓ is used to define $\mathcal{J}_C$, and $\mathcal{J}_C$ is less than τ units in length). Hence, in these cases we need only calculate P(M|M-1), K(M), and P(M|M) for M = 1,2,...,N-1 once. These calculations are then used by all L Kalman filters, which greatly reduces the computational burden.

A flowchart for implementing the L Kalman *filters*, when $\mathcal{S}$ *is time-invariant or slowly time-*varying, is depicted in Figure 4. Because the outer loop varies M and the inner loop varies ℓ, outputs P(M|M) and $\hat{\underline{x}}_\ell(M|M)$ are generated in multiplexed ordering. This can be verified by listing the sequence of outputs from this flowchart and comparing them with the right-hand side of Eq. (81) to see that they are identical.

5. Conclusions

In this paper we have derived the minimum-variance state estimator for uniform causal functional equations. These equations are useful for modeling layered media systems which are described by the lossless wave equation and boundary conditions. Causal functional equations, though continuous-time in nature, bear a strong resemblance to discret-time equations. In fact, we have shown that the solution, $\underline{x}(t)$, of uniform causal functional equation (1) can be generated by an uncountable number of discrete-time systems. We have also shown that for any fixed $t' \epsilon \mathcal{J} = [0,\tau)$, $\hat{\underline{x}}(t)$, where $t = t' + M\tau$ ($M = 1,2,\ldots$), is given by the usual discrete-time Kalman filter equations with t' considered the initial starting time. Of course, to obtain $\hat{\underline{x}}(t)$ for all $t \epsilon \mathcal{R}$ we would need an uncountable number of discrete-time Kalman filters; but, imposing a mesh on $\mathcal{J}$ leads to a countable number of Kalman filters which operate in parallel, as depicted in Figure 3. To the best knowledge of the authors, this is the first estimation theory result that has led to a natural form of parallel data processing.

The results of this paper are not merely an end unto themselves. An important problem for layered media systems is to extract reflection coefficients from noisy measurements. This if often referred to as an inverse problem (Refs. 3, 4, 7, 8 and 10), and usually, solutions are given only for noise-free measurements. The reflection coefficients appear in matrix A. By means of the results of this paper, two appraoches can be studied for solving the inverse problem. In the first approach, we augment state equations (one for each unknown parameter) to Eq. (1) and develop an extended minimum-variance estimator for the resulting augmented system. Of course, the augmented state equations must also be causal functional in nature or else, if they were differential equations, our augmented system would be heriditary in nature and a different course of action would have to be taken. In the second approach, we estimate the reflection coefficients by a maximum-likelihood technique. To accomplish this, we must first develop the correct likelihood function for a uniform causal functional equation. Both of these approaches are presently under investigation.

Acknowledgment

The work reported on in this paper was performed at the University of Southern California, Los Angeles, California, under National Science Foundation Grant NSF ENG 74-02297A01, Air Force Office of Scientific Research Grant AFOSR 75-2797, and Chevron Oil Field Research Co. Contract-76. The first author wishes to acknowledge the support of the Hughes Fellowship Program during the course of this study. The authors also wish to acknowledge Dr. Richard Bucy, of the University of Southern California, for suggesting the Ornstein-Uhlenbeck process to them.

Appendix A. Proof of Theorem 5

Let $\underline{x}_1(t)$ be associated with system , when B=0; that is to say, $\underline{x}_1(t)$ satisfies

$$\underline{x}_1(t+\tau) = A\underline{x}_1(t) + \underline{w}(t) \qquad \text{(A-1)}$$

where

$$\underline{x}_1(\sigma) \triangleq \underline{x}(\sigma) \qquad \forall\ \sigma\epsilon\mathcal{J} \qquad \text{(A-2)}$$

Add Eqs. (A-1) and (53), and (A-2) and (54), to show that

$$\underline{x}_1(t+\tau) + \underline{g}(t+\tau) = A[\underline{x}_1(t) + \underline{g}(t)] + B\underline{m}(t) + \underline{w}(t) \qquad \text{(A-3)}$$

and

$$\underline{x}_1(\sigma) + \underline{g}(\sigma) = \underline{x}(\sigma) \qquad \forall\ \sigma\epsilon\mathcal{J} \qquad \text{(A-4)}$$

Comparing Eqs. (A-3) and (1), and (A-4) and (3), we conclude by uniqueness, that

$$\underline{x}(t) = \underline{x}_1(t) + \underline{g}(t) \qquad \text{(A-5)}$$

Next, we define the modified measurement vector $\underline{y}_1(t)$ in Eq. (52). That $\underline{y}_1(t)$ can be expressed solely in terms of $\underline{x}_1(t)$ is apparent, when Eqs. (2) and (A-5) are substituted into Eq. (52); i.e.,

$$\underline{y}_1(t) = H\underline{x}_1(t) + \underline{n}(t) \qquad \text{(A-6)}$$

Observe that Eqs. (A-1) and (A-6) are now of the right form to use Theorem 4 to obtain $\hat{\underline{x}}_1(t) = E\{\underline{x}_1(t)|Y_1(t)\}$, where $Y_1(t) = \{\underline{y}_1(\lambda): 0 \le \lambda \le t, t\epsilon\mathcal{R}\}$.

From estimation theory, we know that

$$\hat{\underline{x}}(t) = E\{\underline{x}(t)|Y(t)\} \qquad \text{(A-7)}$$

Applying (A-7) to (A-5), we find that

$$\hat{\underline{x}}(t) = E\{\underline{x}_1(t)|Y(t)\} + \underline{g}(t) \qquad \text{(A-8)}$$

since $\underline{g}(t)$ is deterministic. We must now prove that

$$E\{\underline{x}_1(t)|Y(t)\} = E\{\underline{x}_1(t)|Y_1(t)\} = \hat{\underline{x}}(t) \qquad \text{(A-9)}$$

in which case Eq. (A-8) reduces to the desired result in Eq. (55).

We shall examine the sigma fields generated by $\underline{y}(t)$ and $\underline{y}_1(t)$. Let $\mathcal{B}_y$ and $\mathcal{B}_{y_1}$ denote the sigma fields associated with $\underline{y}(t)$ and $\underline{y}_1(t)$, respectively; i.e.,

$$\mathcal{B}_y = \{\omega: \underline{y}(t,\omega) \le \underline{\alpha}\} \qquad \underline{\alpha}\epsilon R^s \qquad \text{(A-10)}$$

and

$$\mathcal{B}_{y_1} = \{\omega: \underline{y}_1(t,\omega) \le \underline{\alpha}\} \qquad \underline{\alpha}\epsilon R^s \qquad \text{(A-11)}$$

From Eq. (52), we see that $\mathcal{B}_y$ can be expressed in terms of $\underline{y}_1(t)$, as

$$\mathcal{B}_y = \{\omega: \underline{y}_1(t,\omega) + H\underline{g}(t) \le \underline{\alpha}\} \qquad \underline{\alpha}\epsilon R^s; \qquad \text{(A-12)}$$

hence,

$$\mathcal{B}_y = \{\omega: \underline{y}_1(t,\omega) \le \underline{\alpha} - H\underline{g}(t)\} \qquad \underline{\alpha}\epsilon R^s \qquad \text{(A-13)}$$

If $\underline{g}(t)$ is finite for any fixed t, $\underline{\alpha} - H\underline{g}(t)$ also ranges through R^s. This condition on $\underline{g}(t)$ is satisfied as long as $B\underline{m}(t)$ is finite for any fixed t. Consequently,

$$\mathcal{A}_y = \{\omega: \underline{y}_1(t,\omega) \le \underline{\alpha}_1\} \qquad \underline{\alpha}_1 \epsilon R^s \qquad \text{(A-14)}$$

which proves that $\mathcal{A}_y = \mathcal{A}_{y_1}$.

Since $\underline{y}(t)$ and $\underline{y}_1(t)$ generate the same sigma fields, conditioning with respect to $\underline{y}(t)$ [or $Y(t)$] is equivalent to conditioning with respect to $\underline{y}_1(t)$ [or $Y_1(t)$]; hence, Eq. (A-9) is true, and, as we pointed out above, Eq. (A-8) reduces to the desired result in Eq. (55).

References

1. P. C. Wuenschel, "Seismogram Synthesis Including Multiples and Transmission Coefficients," Geophysics, vol. XXV, February 1960, pp. 106-129.
2. E. A. Robinson, "Basic Equations for Synthetic Seismograms Using the z-Transform Approach," Geophysics, vol. XXXIII, June 1968, pp. 521-523.
3. E. A. Robinson, Multichannel Time Series Analysis with Digital Computer Programs, Holden-Day, San Francisco, 1967 (Chapter 3).
4. J. F. Claerbout, Fundamentals of Geophysical Data Processing, McGraw-Hill, New York, 1976.
5. A. Trorey, "Theoretical Seismograms with Frequency and Depth Dependent Absorption," Geophysics, vol. 27, no. 6, December 1962, pp. 766-785.
6. S. Treitel and E. A. Robinson, "Seismic Wave Propagation in Layered Media in Terms of Communication Theory," Geophysics, vol. 31, pp. 72-32, February 1966.
7. E. A. Robinson, "Dynamic Predictive Deconvolution," Geophysical Prospecting, vol. 23, December 1975.
8. E. A. Robinson and S. Treitel, "The Spectral Function of a Layered System and the Determination of the Waveforms at Depth," Geophysical Prospecting, vol. 25, 1977.
9. E. Szareniec, "Fundamental Functions for Horizontally Stratified Earth," Geophysical Prospecting, vol. 24, pp. 528-548, 1976.
10. H. Wakita, "Direct Estimation of the Vocal Tract Shape by Inverse Filtering of Acoustic Speech Waveforms," IEEE Trans. on Audio and Electroacoustics, vol. AU-21, October 1973, pp. 417-427.
11. N. E. Nahi and J. M. Mendel, "A Time-Domain Approach to Seismogram Synthesis for Layered Media," presented at the 46th Annual International Meeting of the Society of Exploration Geophysicists, Houston, Texas, October 24-28, 1976.
12. J. M. Mendel, N. E. Nahi, L. M. Silverman, and H. D. Washburn, "State Space Models of Lossless Layered Media," presented at 1977 Joint Automatic Control Conference, San Francisco, California, June 1977.
13. A. Papoulis, "Probability, Random Variables, and Stochastic Processes," McGraw-Hill, New York, 1965.
14. J. S. Meditch, "Stochastic Optimal Linear Estimation and Control," McGraw-Hill, New York, 1969.
15. A. H. Jazwinski, "Stochastic Processes and Filtering Theory," Academic Press, New York, 1970.
16. G. E. Uhlenbeck and L. S. Ornstein, "On the Theory of Brownian Motion," Phys. Rev., vol. 36, 1930, pp. 823-841.
17. J. L. Doob, "The Brownian Movement and Stochastic Equations," Ann. Math., vol. 43, 1942, pp. 351-369.
18. E. Parzen, "Stochastic Processes," Holden-Day, San Francisco, 1962.
19. W. Feller, "Introduction to Probability Theory and Its Applications," vol. 2, John Wiley and Sons, New York, 1966.
20. J. M. Mendel, "Discrete Techniques of Parameter Estimation: The Equation Error Formulation," Marcel Dekker, New York, 1973.

List of Figures

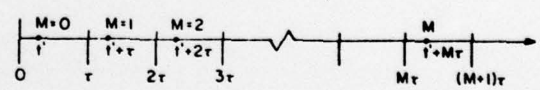

Figure 1

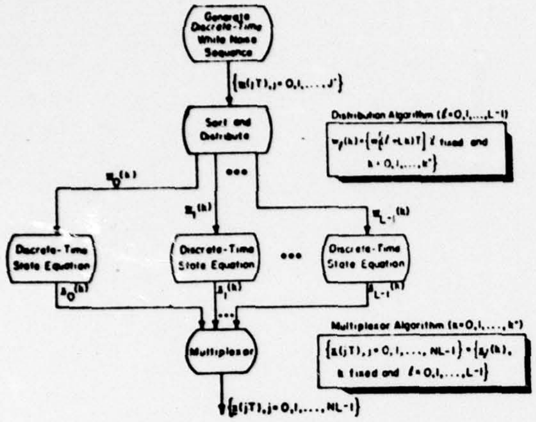

Figure 2

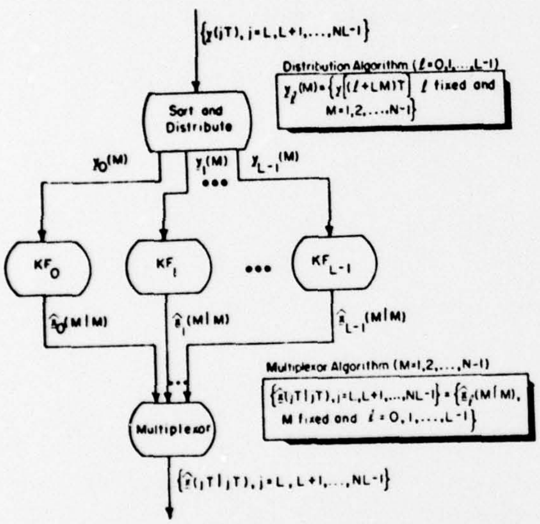

Figure 3

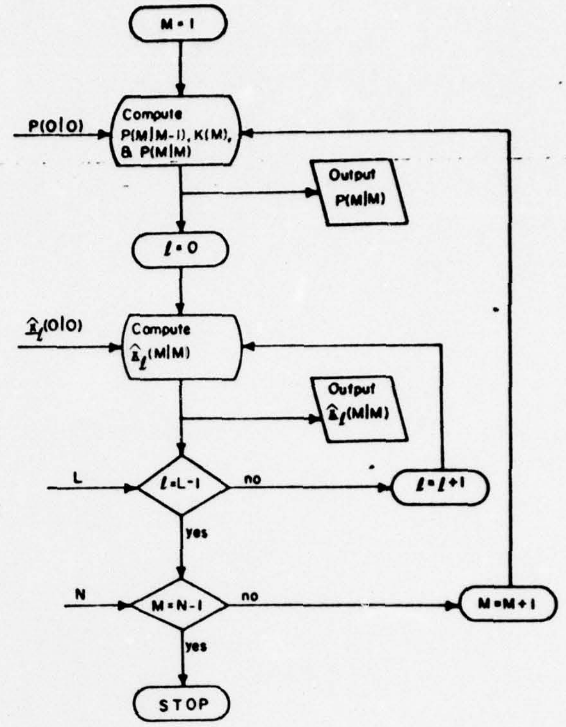

Figure 4